

POWER LAW CONDITION FOR STABILITY OF POISSON HAIL

SERGEY FOSS, TAKIS KONSTANTOPOULOS, AND THOMAS MOUNTFORD

ABSTRACT. We consider the Poisson hail model introduced by Baccelli and Foss. We give a power law condition for the tails (spatial and temporal) of the distribution of jobs to ensure stability as the rate parameter λ tends to zero. We then show that in a weak sense it is optimal.

MSC 2010: 82B44, 82D30, 60K37.

KEYWORDS: Poisson hail, stability, workload, greedy lattice animals

1. INTRODUCTION

The purpose of this article is to loosen conditions for stability in the “Poisson hail” interacting queueing model introduced by [BF]. In this model, there are countably many jobs (identified by countably many points in space-time). Job i requires service τ_i from a subset $B_i \subset \mathbb{Z}^d$. As in the preceding paper, we associate to a job i a (semi-arbitrary) server $x = x(i) \in \mathbb{Z}^d$ who is in some sense central in the group B_i . We suppose that for each site $w \in \mathbb{Z}^d$ the jobs i with $x(i) = w$ arrive according to a Poisson process N_w of rate λ . The jobs i arriving at w will have their subsets B_i and service times τ_i distributed as i.i.d. vectors, so the arrivals at site w may be considered as a marked Poisson process Φ_w . In other words, Φ_w is a Poisson process on $\mathbb{R} \times \mathbb{R}_+ \times 2^{\mathbb{Z}^d}$. Points in its support are typically denoted by (t, τ, B) , the t ’s forming the aforementioned rate- λ Poisson process. The pair (τ, B) is referred to as the mark of the point t . We also assume that the Φ_w are obtained as follows: Let, for each w , $\tilde{\Phi}_w$ be an independent copy of Φ_0 . Then let Φ_w contain all points of the form $(t, \tau, B + w)$ where (t, τ, B) is a point of $\tilde{\Phi}_w$. Thus, the arrival process (including marks) is translation invariant. Physically, we can think of the system as a model of hailstones of cylindrical shape $B \times [0, \tau] \subset \mathbb{Z}^{d+1}$, where τ is the height of the stone and B its base. When a hailstone appears at some point of time at which all sites $w \in B$ are free, it starts melting at rate 1. If there is at least one $w \in B$ occupied by a previously arrived stone, then the current stone will not start melting before all sites in w become free; at the first moment of time this happens, the hailstone starts melting at rate 1. (Only the ground, \mathbb{Z}^d is hot and heat is not transmitted upwards!) At each time t , we let $W(t, x)$ be the

Supported by Swedish Research Council grant 2013-4688.

total work required for x to become empty provided no stones arrive after t . In queueing terms, $W(t, x)$ is a workload. In hailstone terms, $W(t, x)$ is the sum of the heights of all hailstones which contain x in their base and have not been melted yet. Since the superposition of N_w , $w \in \mathbb{Z}^d$, has infinite rate, it follows that within any time interval of positive length there are infinitely many stones arriving. Thus $W(t, \cdot)$ will change infinitely many times in any right neighborhood of t . However, typically, for fixed $x \in \mathbb{Z}^d$, and any $\varepsilon > 0$, $W(t, x)$ depends only on $W(t - \varepsilon, y)$, for y ranging in a finite (but random) number of sites. This is due to the fact that we only have to look at those Φ_w with points (s, τ, B) such that $t - \varepsilon \leq s \leq t$ and $x \in B$.

Fix $x \in \mathbb{Z}^d$ and suppose there is $w \in \mathbb{Z}^d$ such that (t, τ, B) is a point of Φ_w . Then

$$(1) \quad W(t+, x) = \begin{cases} \max_{y \in B} W(t-, y) + \tau, & x \in B \\ W(t-, x), & x \notin B. \end{cases}$$

By convention, we shall assume that $t \mapsto W(t, x)$ is right-continuous: $W(t, x) = W(t+, x)$. On the other hand, if there is no w such that (t, τ, B) is a point of Φ_w with $x \in B$ then $W(s, t)$, $s \geq t$, decreases linearly for a interval of positive length until either it reaches zero or there is job arriving at some $s > t$ at some site w whose base contains x . We have thus completely specified the dynamics of the system. The system considered here differs from that of [KB] in that the latter (i) considers only finitely many sites (\mathbb{Z}^d is replaced by a finite set) but (ii) works for stationary and ergodic arrival processes.

The system is said to be stable if (starting from full vacancy at time 0) the distribution of $W(t, x)$ tight as t varies for fixed x . The central question to be addressed is when is the system stable (for λ sufficiently small). More precisely, for which laws on (τ, B) for jobs arriving at the origin is it the case that there exists $\lambda_0 \in (0, \infty)$ so that the system is stable for all arrival rates $\lambda < \lambda_0$. To avoid trivialities, we assume that B is a finite set, a.s. The founding article [BF] showed that the system was indeed stable provided that there is $c \in (0, \infty)$ so that

$$E[e^{c(\tau + (\text{diam } B)^d)}] < \infty,$$

where $\text{diam } B$ is the diameter of set B , i.e., the maximum of $|x - y|_\infty$ over all $x, y \in B$, and where $|x|_\infty := \max_{1 \leq i \leq d} |x_i|$.

Our purpose in this paper is to slacken this condition to the existence of the $(d + 1 + \varepsilon)$ -th moment for $\tau + \text{diam } B$. We then (easily) show that this condition is (in a certain weak sense) almost optimal. The key idea is to use ideas on laws of large numbers for lattice animals. For this we take as reference the article by James Martin [JM]. In analogy to [KB], one could also ask whether stability is possible for more general arrival processes. This question, however, is outside the scope of our paper as our method explicitly uses the Poissonian assumptions.

Our principal result is

Theorem 1. *Suppose that (τ, B) satisfy, for some $\varepsilon > 0$,*

$$(*) \quad \exists C < \infty \forall x \geq 0 \ P(\tau + \text{diam } B > x) \leq \frac{C}{x^{d+1+\varepsilon}}.$$

Then there exists $\lambda_0 > 0$ so that for job arrival rate $\lambda < \lambda_0$ the system is stable.

Remark. The condition of the theorem is equivalent to the statement that both $E[\tau^{d+1+\varepsilon}]$ and $E[(\text{diam } B)^{d+1+\varepsilon}]$ are finite.

That this result is (in a weak sense) the best possible is shown by

Theorem 2. *For any $d+1 > \varepsilon > 0$, we can find a (spatially homogeneous) job arrival process so that*

$$\exists C > 0 \forall x \geq 0 \ P(\tau + \text{diam } B > x) \leq \frac{C}{x^{d+1-\varepsilon}}$$

and the system is unstable.

Given stability, it is easy to see that starting from complete vacancy (that is no workload at any site), the system converges in distribution to an explicitly describable equilibrium. It is natural to ask whether the system possesses other, not necessarily spatially homogeneous, equilibria. While not definitively answering this we show

Theorem 3. *Under the conditions of Theorem 1, there exists $\lambda_0 > 0$ so that for arrival rate $0 < \lambda < \lambda_0$, the only equilibria for the system that is spatially translation invariant is the limit measure obtained by starting from zero workload.*

We now assemble some observations and techniques from earlier papers, [KB, BF]. Start the system at time $-n$ from full vacancy and consider how the workload $W^n(t, x)$ at time $t \geq -n$ and site $x \in \mathbb{Z}^d$ is obtained.

Definition 1.1. *Let $\Gamma^n(x, t)$ the set of locally constant cadlag paths $\gamma : [u, t] \rightarrow \mathbb{Z}^d$ for some $-n \leq u \leq t$ such that*

- (i) $\gamma(t) = x$,
- (ii) *if $\gamma(s) \neq \gamma(s-)$, then a job arrived at time s requiring service from both servers $\gamma(s)$ and $\gamma(s-)$.*

Associate to such a $\gamma \in \Gamma^n(x, t)$ the score

$$V(\gamma) = \sum_i \tau_i - (t - u),$$

where the sum is over jobs (τ_i, B_i) which arrive at time s_i with $\gamma(s_i) \in B_i$. Based on the way that the workload evolves (see discussion around equation (1)) we obtain that $W^n(t, x) = \sup_{\gamma \in \Gamma^n(x, t)} V(\gamma)$. See Figure 1.

There are three monotonicity properties that the system possesses and which we take into account when analyzing its stability. We start from full vacancy at time $-n$ and consider $W^n(t, x)$ for some $t \geq n$. Then $W^n(t, x)$

will increase if we (i) delay all arrivals between $-n$ and t , or (ii) increase the heights of the stones, or (iii) enlarge their bases.

As discussed and justified in [BF], it will be enough, for the results sought, to consider the case where the sets B for the team of servers required for a job i with $x(i) = 0$ is a cube centred at the origin and we write (for a job arriving at server x in time interval $(m-1, m]$) $R_i^{x,m}$ for the value so that $B = x + [-R_i^{x,m}, R_i^{x,m}]^d$. We will work with time doubly infinite, notwithstanding the fact that we consider the process on $[-n, \infty)$.

The first step is the discretization of the Poisson processes. We consider for $m \in \mathbb{Z}$ and $x \in \mathbb{Z}^d$ the random variables

$$R_{x,m} = \sum_i R_i^{x,m}, \quad T_{x,m} = \sum_i \tau_i^{x,m},$$

the sum taken over all jobs i arriving on the time interval $(m-1, m]$ at site x . It follows from standard methods (see [FKZ]) that

Lemma 4. *If for $\alpha > 0$, there exists $C \in (0, \infty)$ so that $P(\text{diam } B_i > t) < C/t^\alpha$ (resp. $P(\tau_i > t) < C/t^\alpha$), then for all $x, m \exists C' = C'(\lambda)$ so that*

$$P(R_{x,m} > t) < C'/t^\alpha$$

(resp. $P(T_{x,m} > t) < C'/t^\alpha$).

We will deal with the discretized system where at server w at time n a job requiring service time $T_{w,n}$ from each server in the cube $[w-R_{w,n}, w+R_{w,n}]^d$. As discussed in [BF], this discretization is effective in that if we can show stability for the discretized system of jobs then we will have shown stability for the original system: the workload at time m for this system will dominate that arising from the nondiscretized model. It is also as well to note here that we have not given up too much here. In principle if we have multiple w jobs arriving during interval $(m-1, m]$ then we could in principle lose if one job required a long service but only from w while a second job required a very short service from a large cube of servers centred at w . However this will be rare for small λ , where our analysis is most relevant.

2. GREEDY LATTICE ANIMALS (GLA)

As noted, we wish to exploit the celebrated results (see [CGGK]) on greedy lattice animal systems. Recall that a lattice animal of \mathbb{Z}^r is simply a connected subset (when \mathbb{Z}^r is considered as a graph with the standard edge set). We suppose given a collection of i.i.d. positive random variables $\{X(x)\}_{x \in \mathbb{Z}^r}$. We suppose the existence of $\varepsilon > 0$ and $C < \infty$ so that for all $t > 1$,

$$(2) \quad P(X(0) > t) \leq C/t^{r+1+\varepsilon}.$$

(The 1 in the power is unnecessary but it is in this case that we will use our results.) We will then parametrize our system by taking i.i.d. random variables $X^\lambda(x)$ to be equal to $X(x)$ with probability λ and otherwise 0. For a lattice animal $\zeta \subset \mathbb{Z}^r$, its X^λ value (or score) is simply $X^\lambda(\zeta) :=$

$\sum_{x \in \zeta} X^\lambda(x)$. The size of lattice animal ζ , denoted $|\zeta|$, is simply the number of sites in ζ . Thus as λ becomes small the random variables X^λ tend to zero in probability. We fix integer k and $c_1 > 0$ and consider the event

$$A_k^{c_1} = A_k = A_k^\lambda := \{\text{there exists a lattice animal } \zeta \text{ containing } 0 \\ \text{such that } |\zeta| = 2^k \text{ and } X^\lambda(\zeta) \geq c_1 2^k\}.$$

We wish to prove

Proposition 5. *Given any $c_1 > 0$, $\exists \lambda_0 > 0$ and a function $C : [0, \lambda_0) \rightarrow (0, \infty)$ so that $C(\lambda) \rightarrow 0$ as $\lambda \rightarrow 0$ and so that, for $\lambda < \lambda_0$,*

$$P(A_k) \leq \frac{C(\lambda)}{2^{k(1+\varepsilon)}}$$

Remarks. (i) We can use the above to bound the probability that there is a lattice animal of size $u \geq 2^k$ containing the origin whose value is $\geq c_1 u$ for c_1 fixed when λ is small, by considering $\bigcup_{l \geq k} A_l^{c_1/2}$.

(ii) The above formalism will certainly apply to our situation with random variables $R_{x,m} + T_{x,m}$ since if $X(x)$ denotes the random variable at site x for rate $\lambda = 1$ conditioned on there being at least one arrival, then it is easy to see that with rate $\lambda < 1$, $R_{x,n} + T_{x,n}$ is stochastically less than $X^{1-e^{-\lambda}}(x)$.

Proof of Proposition 5. To estimate $P(A_k)$ we split up A_k into 4 parts by fixing v such that

$$q := \frac{1}{1 + \frac{1+\varepsilon}{r}} < v < 1.$$

Then A_k is a subset of the union of

- $A_{k,a} := \{\text{there exists a lattice animal } \zeta \text{ containing } 0 \text{ such that } |\zeta| = 2^k \text{ and}$

$$\sum_{x \in \zeta} X^\lambda(x) I_{X^\lambda(x) \leq 2^{kq}/k^2} \geq \frac{2^k c_1}{10}\}$$

- $A_{k,b} := \{\text{there exists a lattice animal } \zeta \text{ such that}$

$$\sum_{x \in \zeta \cap [-2^k, 2^k]^r} X^\lambda(x) I_{2^{kq}/k^2 < X^\lambda(x) \leq 2^{kv}} \geq \frac{2^k c_1}{10}\}$$

- $A_{k,c} := \{\sum_{x \in [-2^k, 2^k]^r} I_{2^{kv} < X^\lambda(x) \leq 2^{kc_1/2m}} \geq m\}$, where m is such that $m(v(1+r+\varepsilon)-r) > 1+\varepsilon$.

- $A_{k,d} := A_k \setminus (A_{k,a} \cup A_{k,b} \cup A_{k,c})$.

Note that the probability for $A_{k,d}$ is “appropriate” since if $A_k \setminus (A_{k,a} \cup A_{k,b} \cup A_{k,c})$ occurs then there must be a site $x \in [-2^k, 2^k]^r$ with

$$X^\lambda(x) \geq \frac{2^k c_1}{2m},$$

that is, an event of probability that is bounded above by $\frac{K\lambda}{2^{(1+\varepsilon)k}}$. This follows from our assumption (2).

So it remains to bound the probability for $A_{k,a}$, $A_{k,b}$ and $A_{k,c}$.

Bound for $P(A_{k,c})$: The event $A_{k,c}$ is the event that the sum of at most $(2 \times 2^{kr} + 1)^r$ Bernoulli random variables (each with probability being 1 being at most $p_k := \lambda C m^{r+1+\varepsilon} 2^{-kv(r+1+\varepsilon)}$) exceeds m . Thus, the probability of $A_{k,c}$ is less than

$$\left((2^k + 1)^r p_k \right)^m \leq c_m \frac{K\lambda}{2^{(1+\varepsilon)k}}$$

by our choice of m .

Bound for $P(A_{k,b})$: The sum in the definition of $A_{k,b}$ is upper bounded by $2^{kv} \times \text{Bin}(n_k, p_k)$ where $n_k := (2 \times 2^k + 1)^r$ and $p_k = C\lambda \frac{k^{2(r+1+\varepsilon)}}{2^{kr}}$. Thus

$$P(A_{k,b}) \leq P\left(\text{Bin}(n_k, p_k) \geq \frac{c_1}{10} 2^{k(1-v)}\right)$$

and we obtain a good bound by standard bounds on binomial random variables.

Bound for $P(A_{k,a})$:

We repeat the argument given in [JM] (or [CGGK]). We recall Lemma 2.1 which states that for any lattice animal ζ of size n containing the origin and any $1 \leq l \leq n$ we can find a sequence $0 = u_0, u_1, \dots, u_h$ for h the integer part of $2n/l$ and $|u_i - u_{i+1}|_\infty = 1 \forall i$ so that

$$\zeta \subset \cup_{i=0}^h B(lu_i, 2l)$$

where $B(x, u)$ is the L^∞ ball of radius u centred at x . From this it is immediate that for given l there are at most $9^{r2n/l}$ such $2l$ ball coverings.

We use this result with $n = 2^k$. We consider l of “scale” 2^i with $2^i \leq 2^{kq}/k^2$. Let i_0 be the maximal such value. For given i , we choose the value $l = l(i)$ to equal the integer part of $\lambda^{-1/2r} 2^{i/q}$. With this value the probability that a (L^∞) ball of radius $2l$ contains a site having an X^λ value is small for λ small but not (in principle) negligible. From this it is easily seen that given a sequence u_0, u_1, \dots, u_h satisfying the above (and therefore given an l covering), the probability that

the number of sites within the covering having value at least 2^i is at least $2(2^k/l)c_1$

is bounded above by $2^{-r2 \cdot 2^k/l}$ for λ small. Thus we see that outside probability $9^{hr} 2^{-2 \cdot 2^k r/l}$, this bound will hold for all l coverings. Summing over i such that $2^i \leq 2^{kq}/k^2$ we have that outside probability $\sum_{2^i \leq 2^{kq}/k^2} \left(\frac{1}{2}\right)^{2 \cdot 2^k r/l(i)} \leq 2\left(\frac{1}{2}\right)^{2 \cdot 2^k r/l(i_0)} \leq 2\left(\frac{1}{2}\right)^{c(\lambda k^{2/q})}$ for each such i and for each corresponding $l(i)$ covering, the number the number of values of X^λ in the covering sites having value at least 2^i is at most $2(2^k/l)c_1$. Thus (outside of probability $2\left(\frac{1}{2}\right)^{c(\lambda k^{2/q})}$ for some universal c) we have any lattice animal of size 2^k , ζ

$$\sum_{x \in \zeta} X^\lambda(x) I_{X^\lambda(x) \leq \frac{2^{k(r+1+\varepsilon)/d}}{k^2}} \leq \sum_{i \leq i_0} 2(2^k/l(i))c_1 2^{i+1} \leq \sum_{i \leq i_0} 42^k \lambda^{1/2r} 2^{-i(1+\varepsilon)/r}$$

which is bounded by $Constant(\varepsilon)2^k\lambda^{(1+\varepsilon)/2r}$. The conclusion follows for large k .

Thus we have shown the proposition. \square

Corollary 6. *Define*

$$B_u^c(x) := \{ \text{there exists a lattice animal } \zeta \text{ containing } x \\ \text{such that } |\zeta| \geq u \text{ and } X^\lambda(\zeta) \geq c_1|\zeta| \}.$$

For $c_1 < 1$ fixed, there exists a constant $\lambda_1 = \lambda_1(c_1)$ and a function H defined on $[0, \lambda_1)$ tending to zero as λ tends to zero, so that for all $0 < \lambda < \lambda_1$,

$$P\left(\bigcup_{x \in [-R, R]^r} B_u^{c_1}(x)\right) \leq \begin{cases} \frac{H(\lambda)}{(u+1)^{1+\varepsilon}}, & u \geq R \\ \frac{H(\lambda)R^r}{(u+1)^{r+1+\varepsilon}}, & u \leq R. \end{cases}$$

Proof. We treat the case $u \geq R$ only as that for $u \leq R$ is essentially the same. Fix $c_2 < c_1$ let $N = \sum_{x \in [-R, R]^r} I_{B_u^{c_2}}(x)$.

We have by Proposition 5, $E(N) \leq \frac{C(\lambda)(2R+1)^r}{2^{k(1+\varepsilon)}}$.

But if there exists an $x \in [-R, R]^r$ for which $B_u^{c_1}(x)$ occurs then $B_u^{c_2}(y)$ occurs for all $|y - x|_\infty \leq (\frac{c_1 - c_2}{rc_2})R$, so $N \geq c_3 R^r$ for some universal c_3 . The result now follows from Markov's inequality and our bound on $E(N)$. \square

3. CLUSTER FORMATION AND THEIR PROPERTIES

In this section we construct clusters for our Poisson hail corresponding to integer intervals $(m-1, m]$. The clusters themselves will follow a clustering procedure of [BF] and will depend only on the random variables $\{R_{x,m}\}_x$. Our departure will consist in the temporal (or workload) variable we associate to each cluster. Our clusters will have the property that if $C \subset \mathbb{Z}^d$ is a cluster and $\gamma : (m-1, m] \rightarrow \mathbb{Z}^d$ is a path satisfying property (ii) of Definition 1.1, then

$$(3) \quad \gamma(m) \in C \Rightarrow \gamma(s) \in C \quad \forall s \in (m-1, m].$$

Recall we discretized time by identifying with m all tasks for site x arriving in $(m-1, m]$ with a single task of "radius"

$$R_{x,m} \equiv \sum R_i^{x,m},$$

summed over all tasks arriving at x in time interval $[m-1, m]$. We denote by $t_i^{x,m}$ the times of the arrivals, i.e., the points of the Poisson process N_x in the interval $(m-1, m]$. The indices i are coordinated so that for site x a job arrives at time $t_i^{x,m}$ requiring $\tau_i^{x,m}$ units of service from servers in $x + [-R_i^{x,m}, R_i^{x,m}]^d$.

Recall Lemma 4

$$P(R_{x,n} \geq u) \leq \frac{C}{(u+1)^{d+1+\varepsilon}}.$$

The clusters for "time" m are quite simply the connected components of the union of cubes $D_{y,m}$ which are the L^∞ balls centred at y of radius $R_{y,m}$. Thus the cluster $C(x,m)$ containing x is the union of cubes $D_{y,m}$ (centred at y and having radius $R_{y,m}$) having property $D_{y,m} \in C(x,n) \Rightarrow \exists K \ y = z_0, z_1, \dots, z_K = x$ so that $\forall i \ D_{z_i,n} \cap D_{z_{i+1},n} \neq \emptyset$. Let $D(x,m)$ be the diameter of the cluster $C(x,m)$.

It is clear that these clusters have the property (3) above.

A priori it is not clear that even with very small rate λ the clusters will be a.s. finite. However the preceding section enables us to prove

Lemma 7. *There exists a function $K(\lambda)$ tending to zero as λ tends to zero so that for λ sufficiently small and all $z \geq 1$,*

$$P(D(0,n) \geq z) \leq \frac{K(\lambda)}{z^{1+\varepsilon}}.$$

Proof. Consider the GLA system with random variables $\{Z(x)\}_{x \in \mathbb{Z}^d}$ for

$$Z(x) = R_{x,m} + T_{x,m}.$$

If the diameter of the 0 cluster exceeds z then there must exist L and $0 = x_0, x_1, \dots, x_L$ so that $\forall 1 \leq i \leq L, \ |x_{i-1} - x_i|_\infty \leq R_{x_i,m} + R_{x_{i-1},m}$ and $|x_L| \geq z$. We choose ζ to be the lattice animal $\bigcup_i P(x_{i-1}, x_i)$ where $P(x_{i-1}, x_i)$ is a path connecting x_{i-1} and x_i of length $|x_{i-1} - x_i|_1$. (Here, $|x|_1$ denotes the ℓ^1 norm.) Then ζ has a score of at least $(z + |\zeta|)/4d$. The result follows from Proposition 5 applied to $c_1 < \frac{1}{4d}$.

Arguing as in Corollary 6, we obtain

Corollary 8. *There is a function $C(\lambda)$ tending to zero as $\lambda \rightarrow 0$ so that for λ small, for all L , and for $R \leq L/2$,*

$$P(\exists x \in [-R, R]^d \text{ with } D(x,m) \geq L) \leq \frac{C(\lambda)}{L^{1+\varepsilon}}.$$

We now consider the "time", $T(x,m)$, associated with the cluster $C(x,m)$. This definition is a little less direct than that for $D(x,n)$: Given $x \in \mathbb{Z}^d$ and integer m (and so given cluster $C(x,m)$), $T(x,m)$ is less than or equal to the maximum value

$$\sum_{i=0}^L \tau_{j(i)}^{x_i, m}$$

over sequences $x_0, x_1, \dots, x_L \in C(x,m)$ and $m \geq t_0 \geq t_1 \dots \geq t_L \geq m-1$ so that $\forall i$ a job arrives at x_i at time $t_i = t_{j(i)}^{x_i, m}$ having work time $\tau_{j(i)}^{x_i, m}$ and

$$\forall 1 \leq i \leq L, \ |x_{i-1} - x_i| \leq R_{j(i)}^{x_i, m} + R_{j(i)}^{x_{i-1}, m}.$$

We remark that under the latter two conditions, if $x_0 \in C(x,m)$ then necessarily the "subsequent" x_i are also in this cluster. We note also that this

definition (which requires more information than the discretized data) ensures for any site in $C(x, m)$, the waiting time accrued during time interval $(m - 1, m]$ is less than or equal to $T(x, n)$.

Lemma 9. *There exists function $K(\lambda)$ which tends to zero as λ tends to zero so that for λ sufficiently small there exists a finite constant K so that for all $z \geq 1$,*

$$P(T(0, m) \geq z) \leq \frac{K(\lambda)}{z^{1+\varepsilon}}.$$

Proof. By the previous lemma we may suppose that $D(0, n) \leq z/100$. Again if $T(0, m)$ takes a value exceeding z then there must exist L and a path $x_0, x_1 \cdots x_L \in C(0, m)$ and times $m \geq t_0 \geq t_1 \cdots \geq t_L \geq m - 1$ so that $\forall i \leq L - 1$, there is a job arrival at x_i at time t_i and

$$\forall 1 \leq i \leq L, \quad |x_{i-1} - x_i| \leq R_{j(i)}^{x_i, m} + R_{j(i-1)}^{x_{i-1}, m}$$

and also $\sum_i T(x_i, t_i) \geq z$.

It is important to note that we do not assume that the x_i are distinct. Indeed it is for this reason that we use that bound involving $R_{j(i)}^{x_i, m}$ rather than $R_{x_i, m}$. However if y is equal to $x_{i_1}, x_{i_2}, \dots, x_{i_r}$, then of course $R_{y, m} \geq \sum_k R_{j(i_k)}^{x_{i_k}, m}$ and equally $T_{y, m} \geq \sum_k \tau_{j(i_k)}^{x_{i_k}, m}$. Thus as before we obtain (with Z as in Lemma 7) that for a lattice animal $\zeta = \cup_i P(x_{i-1}, x_i)$ that the $\text{GLA}(Z)$ score will exceed $(z + |\zeta|)/4d$.

Again we have

Corollary 10. *There exists function $C(\lambda)$ which tends to zero as λ tends to zero so that for so that for all L and for $R \leq L/2$,*

$$P(\exists x \in [-R, R]^d \text{ with } T(x, m) \geq L) \leq \frac{C(\lambda)}{L^{1+\varepsilon}}.$$

Remark. An alternative method for deriving the results of this section might possibly be based on the results of [BRS].

4. WORKLOAD BOUNDS AND STABILITY

We now apply the foregoing to analyze the workload stability for small values of λ . It is enough to show tightness of $W^n(0, 0)$, that is, the workload at time 0 when the system starts empty at time $-n$. Recall that $W^n(0, 0)$ is obtained as the maximum of scores $V(\gamma)$ where γ ranges in the set of paths $\Gamma^n(0, 0)$. See Definition 1.1.

Due to the monotonicity properties of the system, $W^n(0, 0)$ is readily seen to be bounded above by the quantity $W^{n,D}(0, 0)$ which corresponds to the discretized system and is given by

$$W^{n,D}(0, 0) = \sup_{\gamma} V^D(\gamma)$$

where the supremum is taken over discrete time indexed paths $\gamma : [-r, 0] \rightarrow \mathbb{Z}^d$ for some $0 \leq r \leq n$ satisfying

- (i) $\gamma(0) = 0$,
(ii) for each $-n < i \leq 0$, $\gamma(i-1)$ belongs to cluster $C(\gamma(i), i)$.
The score $V^D(\gamma)$ is given by

$$V^D(\gamma) = \left(\sum_{i=0}^r T(\gamma(-i), -i) \right) - r.$$

We now consider a cube H of length R in $\mathbb{Z}^{d+1} = \mathbb{Z}^d \times \mathbb{Z}$ where the first d co-ordinates are considered as “spatial” and the last one temporal. Accordingly, we write H as $H' \times I$ where I is an interval of length R and H' is a cube of length R in \mathbb{Z}^d . We define the variable

$V(H, u) :=$ number of distinct clusters of value

$$D(x, n) + T(x, n) \geq u \text{ which intersect } H.$$

Clusters are by definition enclosed in a slab $\mathbb{Z}^d \times \{m\}$ for some m and the clusters at different temporal levels m are independent. Thus (after repeated use of Lemma 1 of [BF]) we easily obtain from Corollaries 6 and 10

Proposition 11. *There exists constant K_λ so that, for $u \leq R$, $V(H, u)$ is stochastically less than Poisson of parameter $\frac{K_\lambda R^{d+1}}{(u+1)^{d+1+\varepsilon}}$. For $u \geq R$ it is bounded by a Poisson of parameter $\frac{K_\lambda R}{(u+1)^{1+\varepsilon}}$. Furthermore, as λ tends to zero, K_λ tends to zero.*

This can be applied to the values $V^D(\gamma)$.

Proposition 12. *There exists $\lambda_0 > 0$ and $C < \infty$ so that for $\lambda < \lambda_0$, the probability that there exists an $m \geq n$ so that $V^D(\gamma) > -m/2$, for some $\gamma \in \Gamma_m$, is bounded by C/m^ε .*

Proof. We denote by Γ_m the set of discrete time paths $\gamma : [-m, 0] \rightarrow \mathbb{Z}^d$ satisfying the stipulated conditions. We associate to each path $\gamma(0), \gamma(1), \dots, \gamma(m)$ the score $\sum_{j=0}^{m-1} T(\gamma(j), j) + \sum_{j=0}^{m-1} |\gamma(j) - \gamma(j+1)|_1$. We can identify the path γ with the “lattice animal” consisting of points $(\gamma(i), -i)$ for $i = 0, 1, \dots, m$ together with for each $0 \leq i < m$ the points (y, i) which lie on a path from $(\gamma(i), i)$ to $(\gamma(i+1), i)$ which lies within $\mathbb{Z}^d \times \{i\}$ and has length $|\gamma(i) - \gamma(i+1)|_1$. Thus this lattice animal ζ has size

$$|\zeta| := m + 1 + \sum_{i=1}^m (|\gamma(i) - \gamma(i-1)|_1 - 1)_+ \leq m + \sum_{i=1}^m (|\gamma(i) - \gamma(i-1)|_1).$$

We can associate to points (x, m) in \mathbb{Z}^{d+1} the score equal to its clusters associated time plus diameter. For lattice animals containing the origin $(0, 0)$ we give a score obtained by summing over clusters touched (only once!) the cluster values.

Given Corollaries 8 and 10 and repeated application of [BF, Lemma 1] we can show, arguing as in Section 2, that for any $c_1 > 0$, for $\lambda > 0$ sufficiently

small, the collection of lattice animals starting at $(0, 0)$ having size at least n all have value less than c_1 times their size outside of probability C/n^ε for some C not depending on n . We apply this to the lattice animals given by paths $\gamma : [-m, 0] \rightarrow \mathbb{Z}^d$ satisfying the stipulated conditions with $c_1 < \frac{1}{10d}$. We have that either the size of the associated lattice animal is of size less than $4md + m$, in which case (outside a probability C/m^ε) its score will be less than $m/2$, or it has a size at least $4dm + m$ in which case its score satisfies

$$\sum_{j=0}^{m-1} T(\gamma(j), j) + \sum_{j=0}^{m-1} |\gamma(j) - \gamma(j+1)|_1 \geq \sum_{j=0}^{m-1} |\gamma(j) - \gamma(j+1)|_1$$

which is greater than $(|\zeta| - m)/d$, again an event of probability bounded by C/m^ε . \square

Proof of Theorem 1. From Proposition 12 we have that the discretized system's workload $W^{n,D}(0, 0)$ is tight as n varies. On the other hand, by monotonicity, the limit as $n \rightarrow \infty$ of $W^n(0, 0)$ exists a.s. Tightness ensures that this limit is finite. If we now start the system in full vacancy at time 0 and consider the workload $W(x, n)$ for some $n > 0$, we have that $W(x, n)$ is in distribution equal to $W^n(0, 0)$. Therefore $W(x, n)$ converges in distribution as $n \rightarrow \infty$. \square

Remark. We have actually shown something stronger than tightness: namely, that, starting with an initially empty system, the workload profile at time t converges in distribution, as $t \rightarrow \infty$, to some distribution which we will denote by μ . Standard arguments show that μ is an invariant measure: if we start with $W(0, \cdot)$ distributed according to μ then $W(t, \cdot)$ also has distribution μ . Since $W(t, \cdot)$ is translation invariant in space, this is the case for the limit μ . We have thus proved the existence of an invariant probability measure which is also spatially invariant.

5. NECESSITY AND PROOF OF THEOREM 2

Let for $0 < \varepsilon < d + 1$. We consider the case where the stone heights (job service times) τ satisfy, for t positive integer,

$$P(\tau \geq t) = \frac{1}{t^{d+1-\varepsilon}},$$

and the stone basis B is the cube

$$B = [-\tau, \tau]^d.$$

We consider the number of job arrivals in space time cube $[0, t)^{d+1}$ of duration at least $2t$ for integer t : that is the number of arrivals (B, τ) so that

- (1) $\tau \geq 2t$,
- (2) B is a cube of side length $2\tau + 1$ centred at a site in $[0, t)^d$ and

(3) the job arrives at a time in $[0, t)$.

This random variable has expectation $t^\varepsilon \lambda / 2^{d+1-\varepsilon}$ which for fixed λ tends to infinity as t becomes large. In particular for t large this expectation strictly exceeds $\frac{1}{2}$. We fix such a t now.

Obviously this applies to any translation of the cube and the random variables associated to disjoint space time cubes are independent.

We consider the path γ^n in $\Gamma^{nt}(0, 0)$ which is identically the origin for all $s \in [-tn, 0]$. We note that under our assumptions on (B, τ) any job arriving at a site in $[0, t)^d$ having $\tau \geq 2t$ requires service from the origin. Hence path γ^n has value at least

$$\sum_{j=1}^n 2tX_j - nt$$

where X_j is the number of jobs arriving during interval $[-jt, -(j-1)t)$ and satisfying (1) and (2) above. By the law of large numbers $V(\gamma^n)$ tends to infinity a.s. as n tends to infinity. This is enough to establish instability of the workload in this case no matter what the value of $\lambda > 0$ might be.

6. UNIQUENESS

We now briefly address the question of unicity of invariant measures for the workloads when the power law condition holds and when λ is sufficiently small. We know that if condition (*) is satisfied and parameter λ is sufficiently small then the distribution of workloads, μ , obtained by starting the system at time $-n$ with the workloads identically zero and letting $n \rightarrow \infty$ is invariant. The question that naturally arises is whether other equilibria for the workload, under Poisson arrival of jobs, is possible.

We consider systems that are stationary under spatial translations and show the following

Theorem 13. *Under the condition (*) above, there exists λ_0 so that if the arrival rate λ is less than λ_0 , and ν is an invariant probability for the system on the space of workloads that is preserved by spatial translation, then $\nu = \mu$.*

In this section the assumption that all jobs require service from cubes of servers is not “without loss of generality” so we remark that we only use the weak “irreducibility” condition that for every neighbour of the origin e there exist sequences

$$0 = x_0, x_1, \dots, x_r = e$$

and bases

$$B_1, B_2, \dots, B_r$$

so that $\forall i \ x_{i-1}, x_i \in B_i$ and jobs B_i occur with strictly positive probability.

To show this it suffices to show that for such a measure ν and any bounded cylinder function h , we have

$$\int h(\eta) \nu(d\eta) = \int h(\eta) \mu(d\eta).$$

Given that ν is supposed invariant, this is equivalent to

$$E^\nu[h(W_n)] = \int h(\eta) \mu(d\eta).$$

for any n (and in particular for n large). Given this and our construction of measure μ it will be enough to show that for $\varepsilon' > 0$ and h as above both fixed

$$\left| E^\nu[h(W_n)] - E^{\tilde{0}}[h(W_n)] \right| < \varepsilon'$$

for n large. This will be our objective in the following.

If ν is an invariant measure as described in the statement of Theorem 13, by the ergodic decomposition [K], we may suppose that it is temporally and spatially ergodic. Thus we have that for every M and every site x a.s.

$$\lim_{t \rightarrow \infty} \frac{1}{t} \int_0^t I_{W_s^x \leq M} ds = \nu(W : W^0 \leq M).$$

Thus for an $\varepsilon > 0$ fixed, we can find an M so large that for (every) $x \in \mathbb{Z}^d$ we have with probability at least $1 - \varepsilon$

$$\forall t > M \quad \frac{1}{t} \int_0^t I_{W_s^x \leq M} ds \geq 1 - \varepsilon.$$

let us call the above event B_M^x . Thus we will have by ergodicity that (increasing M if necessary) that with probability at least $1 - 2\varepsilon$

$$\forall k \quad \frac{1}{(2k+1)^d} \sum_{|x| \leq k} I_{B_M^x} > 1 - 2\varepsilon.$$

We now note that at time 0, say, the existence of a large workload V at a site 0, say, implies that with reasonable probability the workload will be of order V for a time of order V in the time interval $[0, V]$.

Proposition 14. *There exists $c_1 \in (0, \infty)$ so that for all V large enough, uniformly over initial workloads $W(0, \cdot)$ with $W(0, 0) > V$, with probability at least c_1 , we have $\forall x \in [-c_1 V, c_1 V]^d$*

$$W(x, t) > V/4 \text{ for } t \in [V/2, 3V/4].$$

Proof. From our “irreducibility” assumptions on the distribution of jobs, it is clear that there exist for each neighbour e of the origin 0 a sequence of jobs with bases B_1, B_2, \dots, B_R so that for each i , $B_i \cap B_{i+1} \neq \emptyset$, $0 \in B_1$ and $e \in B_R$ and the rate at which job with base B_i arrives is strictly positive. Taking R_1 to be the maximum over the R s as neighbour e varies and c to be the minimum over the rates B_i as e and i vary, we obtain that for any x , there exists a “path” B_1, B_2, \dots, B_R so that $R \leq dR_1|x|$, for each i , $B_i \cap B_{i+1} \neq \emptyset$,

$0 \in B_1$ and $x \in B_R$ and the rate at which job B_i arrives is at least c . Thus for every $x \in [-c_1V, c_1V]^d$, the probability that $W(x, V/2) \leq V/2$ is bounded by the probability that a parameter $Vc/2$ Poisson process is less than dc_1VR_1 . The result now follows easily from Poisson tail probabilities. \square

We note that if $V > 4M$ (assuming as we may that $\varepsilon < 1/3$), then $W(x, t) > V/4$ for $t \in [V/2, 3V/4]$ implies that event B_M^x does not occur. This implies that

Proposition 15. *If for some x with $|x| \leq KV$ we have $W(x, 0) \geq V > 4M$, then with probability at least c_1 ,*

$$\frac{1}{(2Kn+1)^d} \sum_{|y| \leq KN} I_{B_M^y} \leq 1 - c_1/(2K)^d < 1 - 2\varepsilon$$

for ε fixed small enough.

This yields the simple corollary

Corollary 16. *For M and ε as above, let $A(V, K)$ be the event that $W(x, 0) > t$ for some $t > V$ and some $|x| \leq Kt$, then under measure ν , the probability that $W(0, \cdot) \in A(V, K)$ is less than ε/c_1 provided $c_1/(2K)^d > 2\varepsilon$.*

From this result our claim is straightforward.

REFERENCES

- [BF] (2011) Baccelli, F. and Foss, S. Poisson hail on a hot ground. *J. Appl. Prob.*, **48A**, 343-366.
- [BRS] (1999) Blaszczyzn, B., Rau, C. and Schmidt, V. Bounds for clump size characteristics in the Boolean model. *Adv. Appl. Probab.* **31**, 910-928.
- [CGGK] (1993) Cox, J.T., Gandolfi, A., Griffin, P., Kesten, H. Greedy lattice animals I: upper bound. *Annals of Probability* **13**, no. 4, 1151-1169.
- [D] (1984) Durrett, R. Oriented percolation in two dimensions. *Ann. Probab.* **12**, no. 4, 999-1040.
- [BD] (1993) Biane, P. and Durrett, R. *Lectures on Probability Theory*. Lectures from the Twenty-third Saint-Flour Summer School held August 18-September 4, 1993. Edited by P. Bernard. Lecture Notes in Mathematics, 1608. Springer-Verlag, Berlin, 1995. vi+210 pp.
- [FKZ] (2011) Foss, S., Korshunov, D. and Zachary, S. *An Introduction to Heavy-Tailed and Subexponential Distributions*. Springer-Verlag, Berlin. vi+123 pp.
- [KB] (1999) Konstantopoulos, and Baccelli, F. On the cut-off phenomenon in some queueing systems. *J. Appl. Prob.* **28**, 683-694.
- [JM] (2002) Martin, J. Linear growth for greedy lattice animals. *Stoch. Proc. Appl.*, **98**, no. 1, 43-66.
- [K] (1985) Krengel, U. *Ergodic theorems*. de Gruyter Studies in Mathematics **6**. Walter de Gruyter & Co., Berlin., viii+357 pp.

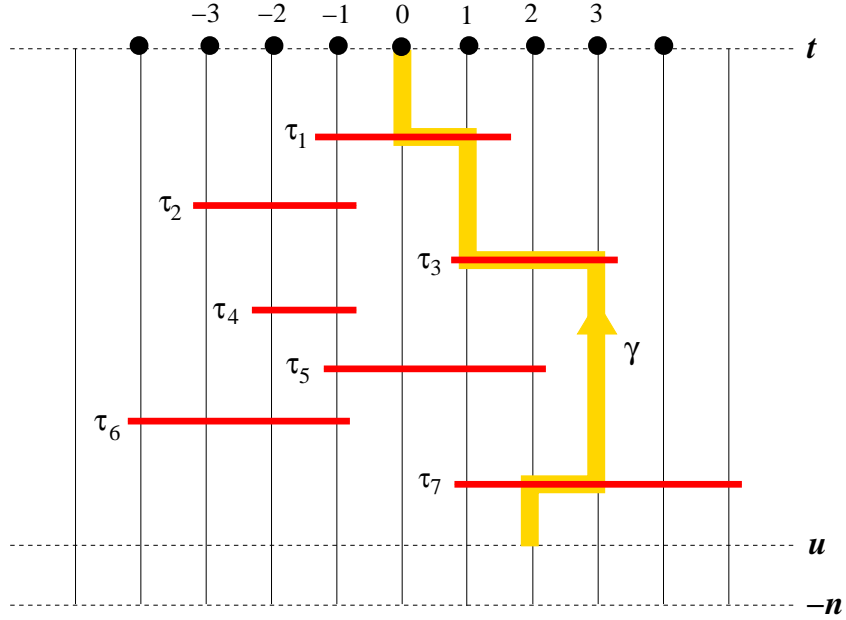


FIGURE 1. Graphical representation of the part of the process responsible for the computation of the profile $W(t, \cdot)$ when the system starts with $W(-n, \cdot) \equiv 0$. Consider the evaluation of $W(t, 0)$ at site $x = 0$. Horizontal intervals represent hailstone (job) arrivals at heights τ_i . Only those arrivals which can potentially influence $W(t, 0)$ are shown. Consider a path γ as indicated, from $(u, 2)$ to $(t, 0)$. Its score is $V(\gamma) = \tau_1 + \tau_3 + \tau_7 - (t - u)$. $W(t, 0)$ is the maximum of these scores over all such paths starting from some (u, y) and ending at $(t, 0)$.

SERGEY FOSS: SCHOOL OF MATHEMATICAL AND COMPUTER SCIENCES, HERIOT-WATT UNIVERSITY, EDINBURGH, EH14 4AS, UK
E-mail address: S.Foss@hw.ac.uk

TAKIS KONSTANTOPOULOS: DEPARTMENT OF MATHEMATICS, UPPSALA UNIVERSITY, P.O. BOX 480, 751 06 UPPSALA, SWEDEN
E-mail address: takiskonst@gmail.com

THOMAS MOUNTFORD: ECOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE, INSTITUT DE MATHÉMATIQUES, STATION 8, 1015 LAUSANNE, SWITZERLAND
E-mail address: thomas.mountford@epfl.ch